Towards an improved understanding of the relative scintillation efficiency of nuclear recoils in liquid xenon

A. Manalaysay[*]

Physik-Institut, Universität Zürich, Zürich, 8057, Switzerland

(Dated: July 23, 2010)

Liquid xenon (LXe) particle detectors are a powerful technology in the field of dark matter direct detection, having shown impressive results in recent years and holding strong possibility for leading the field in sensitivity to galactic weakly interacting massive particles (WIMPs) in the future. The search for WIMPs requires the capability to detect the recoiling nuclei that result when these particles interact with normal matter. In order to make meaningful statements about an observed signal, or lack thereof, the energy scale of recoiling nuclei in LXe must be known. Our understanding of this energy scale is contained in a quantity called the relative scintillation efficiency of nuclear recoils, or $\mathcal{L}_{\rm eff}$, and has been studied extensively in the literature, producing seemingly contradictory results. I examine all the measurements of $\mathcal{L}_{\rm eff}$ that exist, both direct and indirect, and extract the energy dependent behavior that is statistically consistent globally with all values. Additionally, I examine the measurements covering low energies ($\lesssim 10\,{\rm keV}$, where the largest disagreements exist) and attempt to diagnose the systematic effects that have led to the observed inconsistencies. I show that virtually all major disparity arises due to efficiency roll-off of the detectors at the low energies, and, when taking this into account, find that the observed behavior of $\mathcal{L}_{\rm eff}$ supports a slowly and smoothly decreasing value with decreasing energy. Finally, I discuss the prospects for future measurements, and derive a practical limit to what can be achieved.

I. INTRODUCTION

A significant amount of attention has been placed in recent years on the use of liquid xenon (LXe) as a medium for particle detection. In particular, LXe particle detectors have stirred much interest in the field of dark matter direct detection [1–3], experiments that search for Weakly Interacting Massive Particles (WIMPs) leftover from the early universe. In such studies, the interaction of WIMPs in the detector would result in recoiling Xe nuclei of typically less than 100 keV [4].

As a particle loses energy in LXe, it leaves behind a track of ionized and electrically excited Xe atoms (collectively referred to as electronic excitation). Some free electrons liberated by this process can be drifted away from the interaction site and collected for measurement. Additionally, the remaining electrons may recombine with Xe ions, and together with the excited Xe atoms will de-excite in a process that produces 178 nm scintillation photons [5]. These photons can be detected by photomultiplier tubes (PMTs) and converted to an equivalent energy by a quantity called the light yield, which gives the number of photoelectrons (pe) expected per unit of deposited energy.

It is difficult to predict the light yield in absolute terms as it depends on a number of factors, including the energy and identity of the particle. $\mathcal{L}_{\rm eff}$, or the relative scintillation efficiency of nuclear recoils, is defined as the light yield of nuclear recoils relative to that of 122 keV gamma rays from $^{57}{\rm Co}$. The light yield depends on the electronic stopping power (dE/dx), which depends on the energy of the particle, and therefore $\mathcal{L}_{\rm eff}$ is an energy dependent quantity. The deviation of $\mathcal{L}_{\rm eff}$ from unity has several causes. First, while recoiling electrons lose virtually all of their energy through electronic excitation, a significant amount of energy lost by recoiling nuclei goes into heat, and can therefore never be collected as ionization or scintillation [6]. Second, some of the energy going into electronic excitation can be lost by bi-excitonic quenching which removes a fraction of the scintillation photons. The amount of energy lost in this manner depends on the exciton density in the recoil track, which differs greatly for electronic and nuclear recoils [7]. Third, a fraction of the ionized electrons are carried away from the recoil track by thermal motion, preventing them from recombining with ions to produce scintillation photons, and results in a reduced scintillation signal [8]. The fraction of lost electrons depends on the ionization density of the recoil track, again differing for electronic and nuclear recoils. Fourth, the scintillation yield of electronic recoils is also energy dependent, and therefore the energy scale obtained from a calibration at 122 keV is not necessarily valid for electronic recoils lower energies [9]. Fifth, the ratio of excited xenon atoms to ionized xenon atoms ($N_{\rm ex}/N_{\rm ion}$) differs greatly for electronic and nuclear recoils [10].

Measurements of $\mathcal{L}_{\rm eff}$ are performed in two manners, one direct, the other indirect. The direct method uses neutron spectroscopy; a beam of monoenergetic neutrons is incident upon a LXe detector where it scatters, and some of the neutrons scattering under an angle, θ, are tagged by an organic scintillator capable of neutron-

[*]*Electronic address*: aaronm@physik.uzh.ch

gamma discrimination based on pulse shape discrimination (PSD). The true energy of the recoil is given by the kinematics alone,

$$E_r \approx 2E_n \frac{m_n M_N}{(m_n + M_N)^2}(1 - \cos\theta), \qquad (1)$$

where E_r is the energy of the recoiling Xe nucleus, E_n is the energy of the incoming neutron, and m_n and M_N are the masses of the neutron and nucleus, respectively. The approximation is valid when $M_N \gg m_n$ and $E_n \ll m_n c^2$. When the tagging scintillator is placed at a certain angle, one therefore knows the recoil energy and measures the LXe response to those events. The indirect method uses a continuum-energy source of neutrons and compares the resulting spectrum observed in the LXe with a predicted spectrum from a Monte Carlo (MC) simulation. The behavior of $\mathcal{L}_{\mathrm{eff}}$ is then determined by an iterative best-fitting procedure, each time scaling the MC spectrum with a different $\mathcal{L}_{\mathrm{eff}}$ model until the best fit is achieved with the actual data.

Though many direct measurements of $\mathcal{L}_{\mathrm{eff}}$ exist in the literature [11–17], the LXe WIMP search experiments have become increasingly focused upon low-energy interactions, in regions of energy where only three of these measurements provide coverage (below 10 keV) [15–17]. This push to lower and lower energies is due to the fact that WIMPs are expected to produce an exponentially decreasing differential spectrum as a function of energy, and therefore a small improvement in energy threshold yields a large improvement in overall WIMP sensitivity. This improvement in sensitivity comes at a cost, however, because the nuclear recoil energy threshold of a LXe detector can be known only insofar as $\mathcal{L}_{\mathrm{eff}}$ is known. Were the three measurements of $\mathcal{L}_{\mathrm{eff}}$ below 10 keV consistent and precise, no discussion of this sort would be necessary. They are, instead, not consistent, and it is this uncertainty of $\mathcal{L}_{\mathrm{eff}}$ at low energies that provides the largest systematic uncertainty to the results reported in the LXe WIMP searches.

I attack the problem here in two ways. Section II presents a "fair" approach, considering every direct measurement to be equal in validity. Results are weighted only with their reported error bars and the global $\mathcal{L}_{\mathrm{eff}}$ behavior that is statistically consistent with all measurements is extracted. Section III critiques the three direct measurements in the energy region most relevant to WIMP searches, while section IV examines the two indirect measurements. I must point out that I am a co-author on two of these results, one direct [16], one indirect [18], so I run the risk of giving a biased critique. However, fully aware of this risk, I claim no intentional bias, and have made every effort to prevent possible unintentional proclivity.
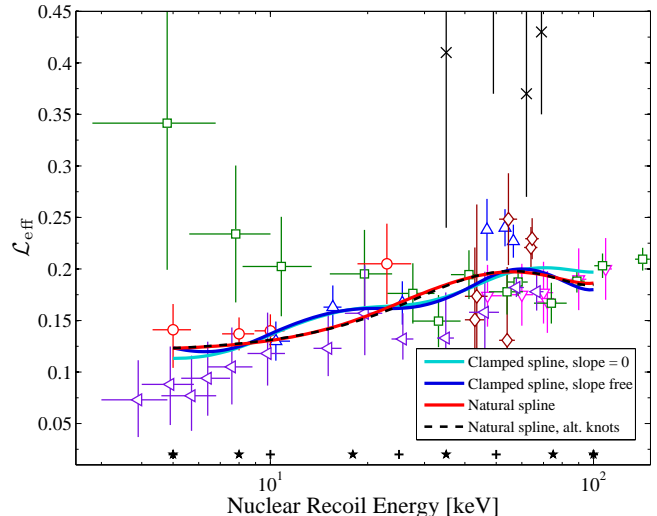

FIG. 1: Measured $\mathcal{L}_{\mathrm{eff}}$ values as a function of Xe nuclear recoil energy. Symbols correspond to ($\triangleleft$)–Manzur *et al.* [17]; ($\circ$)–Aprile *et al.* (2009) [16]; ($\square$)–Chepel *et al.* [15]; ($\triangle$)–Aprile *et al.* (2005) [14]; ($\diamond$)–Akimov *et al.* [13]; ($\times$)–Bernabei et al. [12]; ($\triangledown$)–Arneodo *et al.* [11]. A best fit is performed according to four different methods (see text), three using the spline knots placed at energies indicated by a black '+', and one with alternate spline knots indicated by a black '$\star$'.

II. GLOBAL BEHAVIOR OF $\mathcal{L}_{\mathrm{eff}}$

In an effort to reconcile the various $\mathcal{L}_{\mathrm{eff}}$ measurements, a global fit is performed here to all reported direct measurements. Such a treatment is the most unbiased approach possible, with each measured value weighted by its reported total uncertainty. $\mathcal{L}_{\mathrm{eff}}$ is parameterized as a function of recoil energy by a cubic spline interpolation with spline knots placed at fixed recoil energies. The spline interpolates between knots using cubic polynomials, requiring continuity of the function, and its first and second derivatives. The free parameters of the fit are taken as the value of $\mathcal{L}_{\mathrm{eff}}$ at each spline knot. The energies of the spline knots, 5, 10, 25, 50, and 100 keV, are chosen to be at locations where there are at least three literature measurements in close proximity, with the exception of the knot at 100 keV as only two measurements exist in this range.

The cubic spline is chosen as a parameterization because it permits a large degree of variability in $\mathcal{L}_{\mathrm{eff}}$, and does not assume *a priori* any particular functional shape. In order to calculate a cubic spline, boundary conditions must be chosen which govern the behavior of the interpolation at the endpoints. A *natural cubic spline* forces the interpolation function, f, to have null second derivatives at the endpoints: $f''(x_{\mathrm{first}}) = f''(x_{\mathrm{last}}) = 0$, where x_{first} is the first spline knot, and x_{last} is the last spline knot. A *clamped cubic spline* fixes, instead, the first derivatives at the endpoints: $f'(x_{\mathrm{first}}) = a$, $f'(x_{\mathrm{last}}) = b$, where a and b are arbitrary values. In order to account for these pos-
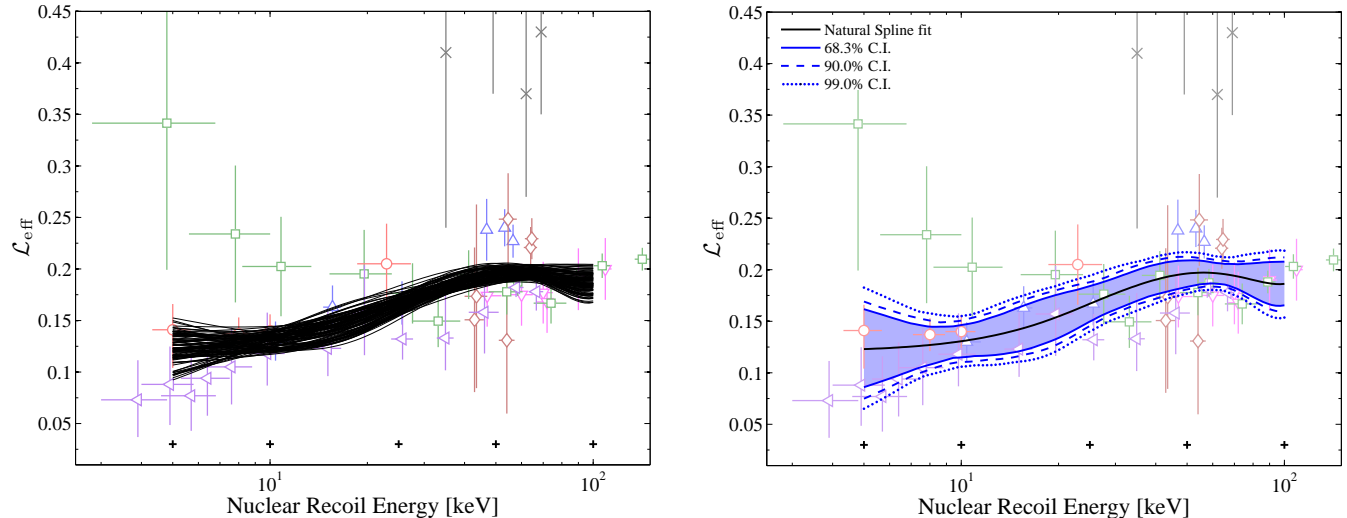
FIG. 2: (*Left*) An example of 100 curves allowed at the 68.3% confidence level. (*Right*) The 68.3%, 90.0%, and 99.0% confidence intervals allowed by all beam measurements in the literature.

sibilities, the fit is performed under three different conditions, seen in Figure 1: (1) a clamped spline with the slope at the endpoints fixed at zero (χ^2/ndf = 1.59); (2) a clamped spline with the slope at the endpoints left as two additional free fit parameters (χ^2/ndf = 1.55); (3) a natural spline (χ^2/ndf = 1.57). The values of $\mathcal{L}_{\text{eff}}$ at the spline knots for these three conditions are shown in Table I. A fourth fit (χ^2/ndf = 1.53) is performed with a natural spline using an alternate set of knot energies (5, 8, 18, 35, 75, and 100 keV). In all cases, the behavior of $\mathcal{L}_{\text{eff}}$ outside of the knot endpoints is extrapolated as a line whose slope is equal to the slope at that endpoint.

TABLE I: The best fit values of the spline fits displayed in Figure 1. Results are shown for a natural spline ("nat."), clamped spline with slope fixed at 0 ("clamp,0"), and clamped spline with endpoint slopes left as free parameters ("clamp,f").

E (keV)	$\mathcal{L}_{\text{eff}}$ (nat.)	$\mathcal{L}_{\text{eff}}$ (clamp,0)	$\mathcal{L}_{\text{eff}}$ (clamp,f)
5	0.123	0.113	0.124
10	0.131	0.136	0.137
25	0.167	0.164	0.162
50	0.196	0.192	0.195
100	0.186	0.197	0.180

The four curves in Figure 1 are calculated in order to estimate the uncertainty introduced by different choices of interpolation. For a single spline choice (natural spline), the uncertainty is determined by the MC method. The procedure choses one random $\mathcal{L}_{\text{eff}}$ value for each spline knot, interpolates, and the resulting χ^2 value is calculated. These steps are carried out 5×10^6 times, each time choosing a new set of random numbers. Confidence intervals (C.I.) are determined by those models

giving $\chi^2 < \chi^2_{\text{min}} + Q_\gamma$, where $Q_\gamma = 5.89$, 9.24, and 15.1 for 68.3%, 90.0%, and 99.0% C.I., respectively, for five free parameters [19]. An example of 100 models within the 68.3% confidence interval are shown in Figure 2 (*left*). The full area covered by all allowed models in the MC is shown in Figure 2 (*right*) for the three confidence levels mentioned here. At 5 keV, the 68.3%, 90.0%, and 99.0% C.I.s allow for variation in $\mathcal{L}_{\text{eff}}$ of ±30%, ±38%, and ±48%, respectively. The same C.I.s allow, at 10 keV, variation in $\mathcal{L}_{\text{eff}}$ of ±12%, ±15%, and ±19%, respectively.

III. COMPARISON OF DIRECT MEASUREMENTS

The analysis performed in section II represents the true global trend of all reported beam measurements, and suggests a slowly decreasing value of $\mathcal{L}_{\text{eff}}$ below ~50 keV. However, though this method is fair in that it weights measurements only by their reported error bars, it has no power to make conclusions regarding the true behavior of $\mathcal{L}_{\text{eff}}$ in the lowest energies. Theoretical attempts have been made to understand the energy dependent behavior of $\mathcal{L}_{\text{eff}}$ [7, 17]. However, these studies are based on Lindhard theory [6] of nuclear quenching which is invalid in LXe below 10 keV [20, 21], and we must instead rely on measurements alone. The discrepancy between the three measurements below 10 keV has yet no explanation. Seen in Figure 3, rising $\mathcal{L}_{\text{eff}}$ behavior (for decreasing energies) below ~30 keV is supported by Chepel [15]; the results from Aprile [16] are consistent with a flat, or slowly decreasing $\mathcal{L}_{\text{eff}}$ below 10 keV; Manzur [17] shows a slowly falling $\mathcal{L}_{\text{eff}}$ until ~10 keV, followed by a sharp drop below this energy. To address this issue, a thorough comparison of the various strengths and weaknesses of each measure-

ment is here made. Only after such a comparison can one make any reasonable statements about which of the three behaviors most closely represents reality.
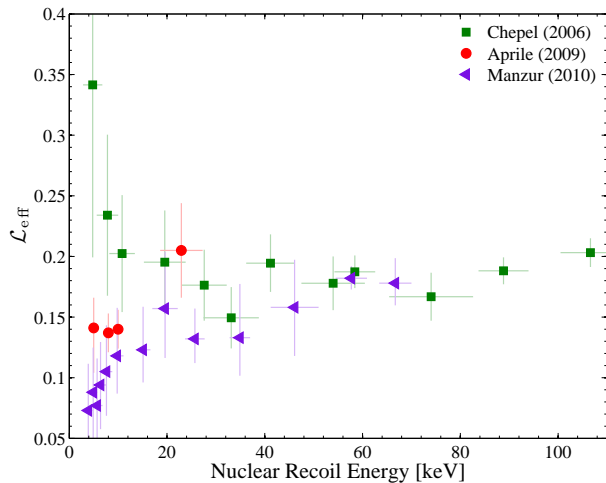

FIG. 3: Results of the three direct $\mathcal{L}_{\mathrm{eff}}$ measurements that extend below 10 keV, shown alone to highlight the disagreement at low energies.

At first glance, measurement of $\mathcal{L}_{\mathrm{eff}}$ in a beam study is simple. By tagging neutrons that scatter only under a chosen angle, a peaked scintillation spectrum in the LXe detector is produced. A calibration of the detector with ^{57}Co (122 keV gamma rays) determines the detector's light yield, L_y, which measures how many photoelectrons (pe) are collected from the photomultiplier tubes (PMTs) for a given energy deposition. The scintillation signal, in pe, from the tagged neutron data can be converted to an energy in 'keVee', or 'keV, electronic recoil equivalent', by scaling the signal with L_y. $\mathcal{L}_{\mathrm{eff}}$ is then simply found by the ratio of the observed electronic recoil equivalent energy to the true recoil energy:

$$\mathcal{L}_{\mathrm{eff}} = \frac{E_{ee}}{E_{nr}}, \qquad (2)$$

where E_{ee} is the observed electronic recoil equivalent energy (based on 122 keV gamma rays), and E_{nr} is the true energy of the nuclear recoil (fixed by the kinematics). This relation holds only for zero applied electric fields, but inclusion of non-zero fields can be done by multiplying the right hand side of equation 2 by S_e/S_n, where $S_{e(n)}$ is the field quenching factor for electronic (nuclear) recoils. Systematic uncertainties, therefore, fall into two categories: those affecting E_{nr}, and those affecting E_{ee}. Table II compares many of the properties of each experiment that can lead to significant systematic uncertainties.

A. Systematics affecting E_{nr}

The origin of systematic uncertainties in E_{nr} can be found by examining equation 1. For neutrons generated by a beam, the resulting neutron energy, E_n, can have a spread due to the thickness of the target. This is because the projectile (deuterons in [15] and [17], protons in [16]) can lose energy in the target before producing a neutron. The neutron energy, and energy spread, are listed in Table II. The fraction of energy lost by the projectile in the target decreases with increasing projectile energy, which explains why the neutron energy of Chepel is more tightly constrained than that of Aprile and Manzur. The energy of the neutrons reported in [17] (2.8 MeV) is valid for forward neutrons produced by 80 keV deuterons incident upon deuterium; however, the thickness of the target is much larger than the penetration depth of the deuterons [23]. This means that 2.8 MeV represents the *maximum* neutron energy, with the average energy lying below this value. Additionally, neutrons produced at other angles (lower energies) can scatter to the forward direction, thus bringing the average neutron energy down even further, to around 2.6 MeV. A lower true neutron energy in equation 1 translates into a lower true recoil energy, which in turn translates through equation 2 to systematically higher $\mathcal{L}_{\mathrm{eff}}$ values. This could partially explain why the Manzur data points appear to be systematically lower (at all energies) than other measurements.

A spread in E_{nr} can also arise due to the finite size of the LXe and tagging detectors. Though the tagging scintillator is placed at a fixed scattering angle, this angle is the *true* scattering angle only for those neutrons scattering in the center of the LXe, and in the tagging scintillator along the axis connecting its center to the LXe's center. In reality, neutrons may scatter anywhere in the volume of both detectors, yielding a range of true scattering angles, and thus, a range of true recoil energies. The degree to which this range is confined depends on the size of both detectors, and on the distance of separation between the two detectors. These parameters are compared in the eighth column of Table II. Also listed is the range of scattering angles allowed when the tagging scintillator is placed at 45°. As the size of the tagging scintillator used by Chepel is not mentioned in their paper, I have here assumed it to be 3 inches (7.6 cm) in diameter, typical of this sort of detector.

Additional systematic effects pertaining to the true recoil energy involve non-peaked backgrounds coming from events with multiple interaction vertices. Such an event can be one in which the neutron scatters multiple times in the LXe ("multiple scatter"), or one in which the neutron scatters once in the LXe and once in any of the detector materials surrounding the active LXe volume ("materials scatter"). Multiple scatters are minimized by reducing the LXe volume (as done by Aprile), while materials scatters are controlled by minimizing the mass of detector materials (as done by Manzur). The detector used by Aprile contained large amounts of PTFE sur-

TABLE II: Comparison of the systemic properties of Chepel [15], Aprile [16], and Manzur [17], the three measurements of $\mathcal{L}_{\text{eff}}$ below 10 keV. Row labels indicate: ('n-gen') process for neutron generation (projectile $\to$ target), D, T, and P stand for deuterium, tritium, and protons, respectively.; ('E_n') reported energy of neutrons and energy spread; ('Collim.?') whether or not the neutron beam was collimated; ('Photo cov.') fraction of LXe volume surface covered by photodetectors; ('LXe dia.') diameter of LXe volume; ('nScint. dia.') diameter of organic tagging scintillator; ('LXe-nScint') distance between LXe volume and tagging scintillator; ('Δ_θ @45°') range of possible scattering angles for a measurement at 45°; ('L_y') light yield from ^{57}Co; ('^{57}Co E-res.') energy resolution of the ^{57}Co peak, as σ/μ; ('Res. acc.?') whether or not the analysis considered the uncertainty in the energy resolution; ('ϵ (5 keV)') the detection efficiency for the measurement of 5 keV nuclear recoils; ('Effic. acc.?') whether or not the analysis accounted for the detection efficiency in LXe.

Study	n-gen	E_n (MeV)	Collim.?	Photo cov.	LXe dia.(cm)	nScint dia.(cm)	LXe-nScint	Δ_θ @45°	L_y (pe/keV)	^{57}Co E-res.	Res. acc.?	ϵ (5 keV)	Effic. acc.?
Chepel (2006)	D→D	6-8±1%	yes	20%	16.5	n/a	1 m	10°[a]	5.5	7.6%	no	~30%[b]	no
Aprile (2009)	P→T	1.0±8%	no	95%	2.5	7.6	50 cm	9°	20	9.0%	no	97%	yes
Manzur (2010)	D→D	2.8[c]	no	56%	5.0	7.6[d]	12 cm[e] 20 cm	46° 26°	11	8.8%	yes	44%	yes

[a] Assuming nScint dia. is 7.6 cm.
[b] Estimated based on L_y.
[c] The energy spread, not reported in [17], is estimated by the manufacturer [22, 23] of the neutron generator used in this study to be ±8%. Additionally, due to the thickness of the target and the angular acceptance, the mean neutron energy is ~2.6 MeV (see text).
[d] Incorrectly reported in [17] as 3.8 cm [24].
[e] The distance between LXe and nScint. varied in this range depending on the available space [24].

rounding the active LXe volume. This was used in order to fill the space between the active LXe and the walls of the much larger chamber that would otherwise be filled with passive LXe ("passive", meaning it is not viewed by PMTs). Both materials, PTFE and passive LXe, can present additional scattering points for neutrons before and after they scatter in the active LXe. However, given the lower mass of nuclei in PTFE (F–19 u, C–12 u) than Xe (131 u), neutrons tend to lose much more energy through scatters in PTFE than in LXe. This lowered energy can then usually be identified as an increased time of flight (ToF) between LXe and tagging scintillator. Nevertheless, the large amounts of PTFE used introduced considerable backgrounds in the results of [16], even after cutting events with large ToF. These backgrounds were estimated by MC simulations and subtracted, but the uncertainty in this background estimate contributed significantly to the total uncertainties reported.

B. Systematics affecting E_{ee}

Systematic processes affecting the observed energy, E_{ee}, are due to the properties of the LXe detector itself. The energy resolution of the detector can cause a true energy peak to appear skewed, because signals on the low energy side of the peak experience proportionally larger Poisson fluctuations than those on the high energy side. This effect generally pulls the measured $\mathcal{L}_{\text{eff}}$ lower than the true value (since low-energy recoils experience proportionally larger fluctuations than the high-energy recoils), and is highly dependent on the observed number of pe. For example, the light yields, L_y, of the detectors used by Chepel, Aprile, and Manzur are 5.5, 20, and 11 pe/keVee, respectively. For nuclear recoils of 5 keV, the energy resolution (assuming Poisson fluctuations on the number of pe and amplification fluctuations of the PMTs) gives a systematic error on the calculated $\mathcal{L}_{\text{eff}}$ in these studies of roughly 40%, 0.1%, and 5%, respectively, in the case that the true value of $\mathcal{L}_{\text{eff}}$ at this energy is 0.123 (here, I have assumed that the spread in E_{nr} is given by the horizontal error bars of the 5 keV point from each measurement in Figure 3). Deviations of the pe fluctuations from Poissonianity can change these errors by factors of a few. Additionally, the effect is worsened if the spread in the true recoil energy, E_{nr}, is large, and also if the true value of $\mathcal{L}_{\text{eff}}$ is changing rapidly within that spread.

Detection Efficiency

A more significant effect results from the total photon detection efficiency of the detector. This total detection efficiency is the combination of the geometrical light collection efficiency (LCE), quantum efficiency (QE) of the PMT photocathodes, collection efficiency of the first dynode, PMT amplification effects, hardware trigger efficiency, and software acceptance. The geometrical LCE, QE, and the hardware trigger efficiency usually provide the dominant contributions. The first efficiency (LCE) varies depending on where in the LXe volume the energy is deposited. Because the spatial distribution from ^{57}Co gamma rays differs from that of neutrons, a systematic shift in the reconstructed $\mathcal{L}_{\text{eff}}$ value can arise in situations where the spatial distribution of the two classes of
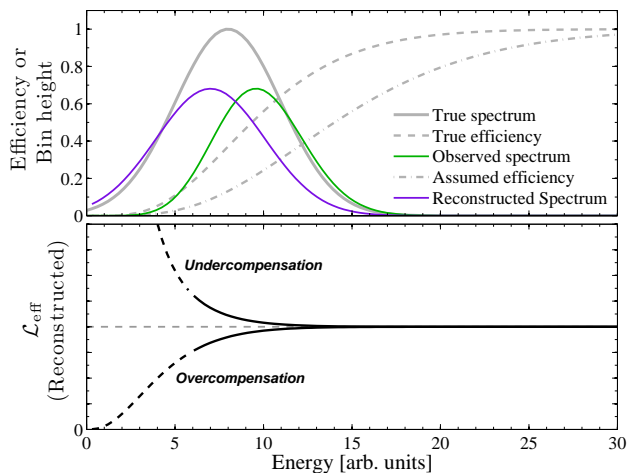
FIG. 4: (*Top*) An example of how detection efficiency affects the position of a peak. The solid grey curve is a hypothetical true peak, but due to a falling trigger efficiency in this region (grey dashed) it is observed as the green curve, whose peak is shifted to higher energies. If the experimenter attempts to compensate the spectrum with an incorrect trigger efficiency (grey dot-dashed), the resulting peak (purple) will be shifted below the true value. (*Bottom*) For the hypothetical situation where the true $\mathcal{L}_{\mathrm{eff}}$ behavior is flat with energy (grey dashed), undercompensation and overcompensation for the total detection efficiency results in a reconstructed $\mathcal{L}_{\mathrm{eff}}$ that rises and falls, respectively, with decreasing energy.

FIG. 5: The reported $\mathcal{L}_{\mathrm{eff}}$ values from Aprile [16] (red circles) and Manzur [17] zero field (dark purple triangles) and $1.5\,\mathrm{kV/cm}$ (light purple triangles) as a function of electronic recoil equivalent energy. Superimposed are the detection efficiencies of each study. The trigger and acquisition used by Manzur differed for single phase (SP) and dual phase (DP).

events differs greatly. For such small detectors as used in these three studies, this effect from spatially-varying light yields is minimal.

At low energies, the total detection efficiency produces a much more significant effect. Due to, for example, Poisson fluctuations on the number of detected photoelectrons, a low energy event will have a probability of being detected that is less than unity. Figure 4 (*top*) shows how such a detection efficiency "roll-off" can affect the position of the observed spectrum. If this detection efficiency is strongly changing in the region of the peak, the observed peak will be shifted to higher energies than the true peak (green curve). As the true recoil energy decreases, the observed peak therefore stays at a relatively constant value. It is clear, then, from equation 2 that such an effect (decreasing E_{nr} with constant E_{ee}) results in a reconstructed $\mathcal{L}_{\mathrm{eff}}$ value that rises at low energies. Figure 4 (*bottom*) shows the resulting $\mathcal{L}_{\mathrm{eff}}$ in the case that this detection efficiency is either unaccounted—or undercompensated for.

Because the roll-off of the detection efficiency is related to the probability to detect a photon, the energy scale at which this roll-off becomes important is related to the light yield, L_y. One therefore expects that the detector used by Chepel, having the lowest value of L_y, to show the strongest effects due to the efficiency roll-off. Though Chepel's detection efficiency is not mentioned in [15] or [25], it can be roughly estimated by scaling Manzur's
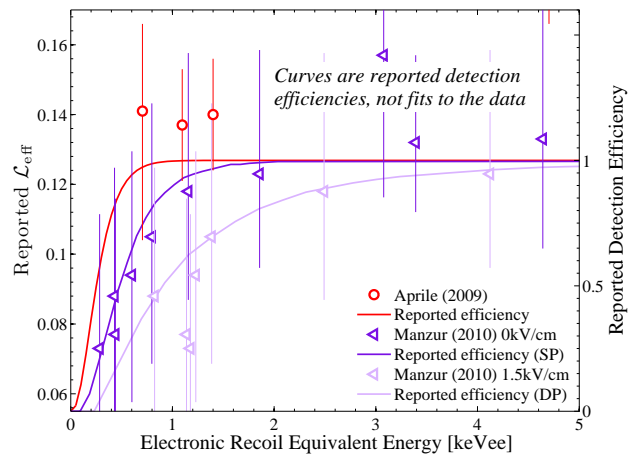
efficiency by L_y. This estimation predicts that the efficiency roll-off becomes significant at around $2\,\mathrm{keVee}$, or for measurements with recoil energies less than $\sim 12\,\mathrm{keV}$. It is precisely at this energy where the $\mathcal{L}_{\mathrm{eff}}$ values of Chepel begin to rise significantly. One can therefore conclude that the rise in $\mathcal{L}_{\mathrm{eff}}$ observed by Chepel *et al.* is non-physical, and is instead an artifact of their detection efficiency.

If, on the other hand, one takes into account the trigger efficiency, as has been done by Aprile and Manzur, the systematic shift of the observed peak to higher energies should be removed. However, a precise treatment of this procedure requires precise knowledge of the detection efficiency. If the estimated detection efficiency is lower than the true efficiency, the reconstructed peak position will again become shifted. Figure 4 (*top*) shows the effect of compensating the observed spectrum with an incorrect estimate of the detection efficiency (purple curve); in particular, with one that is lower than the true efficiency. In this case, the reconstructed peak position is *lower* in energy than the true peak position. Such an effect leads to a measured $\mathcal{L}_{\mathrm{eff}}$ value that is artificially lower than the true value, seen in Figure 4 (*bottom*).

The estimated detection efficiency used by Manzur is reported in [17] as a function of the number of photoelectrons and that of Aprile is reported in [26] as a function of E_{ee}. $\mathcal{L}_{\mathrm{eff}}$ is reported as a function of E_{nr}. In order to make a comparison, the two quantities ($\mathcal{L}_{\mathrm{eff}}$ and detection efficiency) can be independently scaled as a function of E_{ee}. For $\mathcal{L}_{\mathrm{eff}}$, this is done by simply inverting equation 2. Figure 5 shows this comparison for the results reported by Aprile and Manzur. In the range of energies investigated by Aprile *et al.*, the detection efficiency is near unity, dropping down to 97% at the value with the lowest energy ($5\,\mathrm{keV}$ nuclear recoils). For these data,

therefore, systematic effects due to detection efficiency are minimal. In contrast, the values reported by Manzur show a steep drop in $\mathcal{L}_{\text{eff}}$ coinciding *exactly* with their reported efficiency roll-off, so much so that the efficiency estimate almost appears to be a best fit when it is superimposed over the data, shown in Figure 5. It seems unlikely that the laws of physics would produce a major change in $\mathcal{L}_{\text{eff}}$ with exactly the same energy dependence as an important detector pathology inherent in the Manzur measurement. This correlation indicates that the low energy behavior observed by Manzur is almost certainly an artifact of efficiency overcompensation. A valid extrapolation of the low-energy behavior of $\mathcal{L}_{\text{eff}}$ using the results of Manzur must therefore consider only those data with recoil energies greater than 10 keV.

IV. COMPARISON OF INDIRECT MEASUREMENTS

The indirect method of determining $\mathcal{L}_{\text{eff}}$, described in section I, has been performed using data from the XENON10 [18] and ZEPLIN-III [2] detectors. In these studies, $\mathcal{L}_{\text{eff}}$ is parameterized as a function of nuclear recoil energy; the model parameters are then varied until the best fit is achieved between the observed and simulated neutron spectra. As a cubic spline interpolation was used here in section II, so was this method also used to parameterize $\mathcal{L}_{\text{eff}}$ in [18]. The model used in [2], however, is not stated, and appears to be a sigmoid function with likely two or three free parameters. If this is the case, then the model is forced to fall sharply at decreasing energies, regardless of the parameter values. Indeed, the high energy behavior of $\mathcal{L}_{\text{eff}}$ reported in [2] is perfectly flat, which is unsupported by all other measurements.

In section III, comparison of the $\mathcal{L}_{\text{eff}}$ energy dependence with the detection efficiency was extremely illustrative in diagnosing the observed results. Here as well, the same comparison can be made. However, while efficiency undercompensation and overcompensation had clearly predictable effects in direct beam measurements, the effect on indirect measurements is less obvious. Undercompensation of trigger efficiency results in underpopulated histogram bins in the low energies, which will obviously disagree with the simulation. However, the simulated spectrum can be made to fit this deficient measured spectrum, not by adjusting the bin heights according to the efficiency, but by suppressing $\mathcal{L}_{\text{eff}}$ in this range. Doing so pulls the low-energy simulated events below the energy threshold of the fit range. In this way, depopulation of the low energy bins is done not by vertical scaling (efficiency compensation), but by horizontal scaling (energy compensation). The result is a $\mathcal{L}_{\text{eff}}$ curve that is artificially falling.

Figure 6, analogous to Figure 5, shows the reported $\mathcal{L}_{\text{eff}}$ curves from ZEPLIN-III [2] (blue) and XENON10 [18] (red) as a function of electronic recoil equivalent energy with the reported detection efficiency curves super-
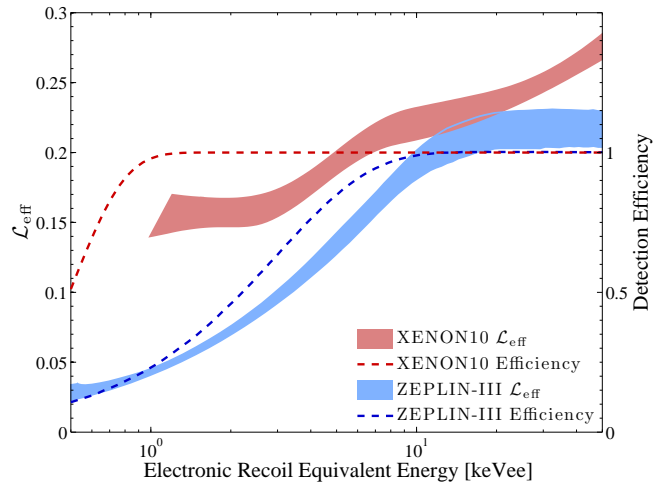

FIG. 6: The reported $\mathcal{L}_{\text{eff}}$ curves by XENON10 and ZEPLIN-III determined using the indirect method with a neutron source having a continuum energy spectrum (AmBe). Superimposed are the energy dependent part of each detectors' efficiencies. The energy independent efficiencies are 93.7% and 59.1% in XENON10 and ZEPLIN-III, respectively.

imposed. Not only do the two $\mathcal{L}_{\text{eff}}$ curves have virtually no overlap, they are made in regions where their efficiencies differ immensely. The energy range of validity allowed by [18] sees an efficiency of no less than 99%. In contrast, the $\mathcal{L}_{\text{eff}}$ curve reported in [2] closely follows the measured efficiency roll-off. This efficiency was not taken into account in [2], and therefore the reported $\mathcal{L}_{\text{eff}}$ behavior is likely an artifact. Indeed, if this represented the true behavior of $\mathcal{L}_{\text{eff}}$, then the beam measurements discussed in section III would see no signal at all in the low energies.

A possible problem of the results found by the indirect technique has recently been suspected [27, 28]. Neutrons from the source can scatter multiple times in LXe. The time between interactions, $\mathcal{O}(10\,\text{ns})$, is similar to the time-scale of the scintillation emission, and therefore the multiple scintillation signals will be observed as a single pulse. However, these multiple scatters are vetoed in time projection chambers (TPCs) like XENON10 and ZEPLIN-III by the observation of multiple ionization signals (separated by tens of μs). In such detectors, however, there generally exist regions of LXe outside the active volume that are insensitive to charge, but nonetheless have a non-negligible optical coupling to the active LXe volume. A neutron that scatters once in the active LXe volume, and once in a charge insensitive region will thus be reconstructed as a single scatter, with a scintillation signal that is anomalously too large since it is the sum of two interactions.

To some extent, these events can be removed from the data by cuts in energy and also the ratio of scintillation to ionization signals. However, the effectiveness of these cuts is difficult to quantify, as is the optical coupling of the charge insensitive LXe regions. Therefore, the spec-

tra from data and simulation will not represent *exactly* the same class of events. But this pathology does not render the indirect measurements invalid. The presence of multiple scatters affects mainly the high energy regions of the observed spectrum, because these are effects that only *add* scintillation light. A discrepancy in the high energies affects mainly the overall vertical offset of the reconstructed $\mathcal{L}_{\mathrm{eff}}$ curve. However, the low energy shape of the result remains robust.

V. TOWARDS AN IMPROVED UNDERSTANDING

The rather disparate landscape of $\mathcal{L}_{\mathrm{eff}}$ measurements, both direct and indirect, casts doubt on our ability to currently resolve the underlying nature of this illusory quantity at low energies. However, in sections III and IV it was shown that detection efficiency roll-off has, so far, *always* produced sizable pathological effects on the determination of $\mathcal{L}_{\mathrm{eff}}$. We can then use this understanding as a lens to focus our view of the landscape. Figure 7 shows the available measurements with two modifications: (1) only measurements in regions where their detection efficiency is greater than 90% are shown; (2) the results of [17] have been corrected for a more reasonable estimate of the neutron energy. Viewed in this way, a much clearer trend emerges: the combined results imply a $\mathcal{L}_{\mathrm{eff}}$ with no sharp changes with energy. At 10 keV, three measurements [14, 16, 17] show remarkable consistency of a value of $\mathcal{L}_{\mathrm{eff}} \simeq 0.135$.
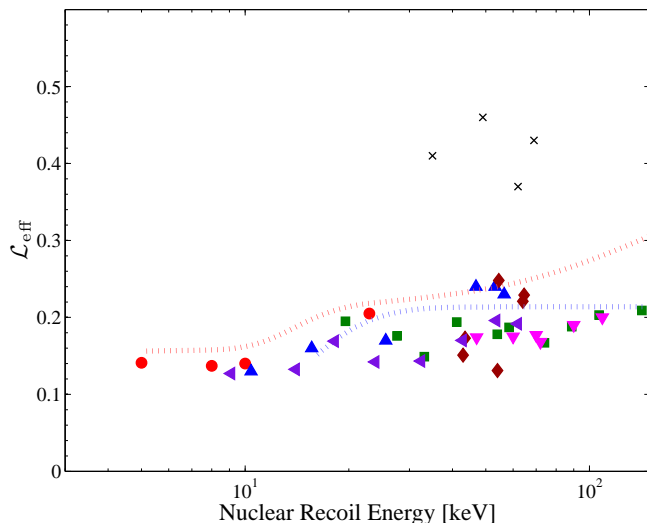

FIG. 7: The landscape of $\mathcal{L}_{\mathrm{eff}}$ measurements, considering only results where the detection efficiency exceeds 90% at the reported energy. Additionally, the results of [17] have been corrected using a more reasonable estimate of the neutron energy. Markers and colors correspond to the same as those in Figures 1 and 6.

The behavior of $\mathcal{L}_{\mathrm{eff}}$ at energies below what are dis-

cussed here is difficult to predict. Despite the appearance of equation 2, $\mathcal{L}_{\mathrm{eff}}$ does not measure the ratio of scintillation response of nuclear to electronic recoils of the same energy. This is because L_y, as it is used in the definition of $\mathcal{L}_{\mathrm{eff}}$, is valid only for electronic recoils at 122 keVee, while it is known that the true electronic recoil L_y grows at low energies [9]. If the true ratio of scintillation response from nuclear to electronic recoils drops at low energies (as it is expected to do), the nonlinearity of the electronic recoil light yield will act to soften that drop.

The need for continued measurement of $\mathcal{L}_{\mathrm{eff}}$ is obvious, but what improvements can be made upon existing measurements? The lessons of section III are clear on this. The spread in the true recoil energy must be minimized. The contribution from the spread in the neutron energy is difficult to minimize, so instead one must focus on building a LXe chamber having a small volume, with the tagging scintillator(s) placed far away. This must be coupled with a minimization of backgrounds due to multiple scatters in surrounding detector materials. Light collection must be maximized, both to reduce effects from efficiency roll-off, but also to reduce uncertainties caused by energy resolution.

It appears that, although efficiency roll-off has so far caused problems, this systematic is one that cannot be completely avoided. The measurements of Aprile *et al.* [16] were performed using a detector having virtually 4π photo coverage of the LXe volume, which is why L_y in that study (20 pe/keVee) was so much larger than in any other similar study. For this reason, the detection efficiency obtained in this study likely represents the practical limit in performance, and therefore any attempts at measurements below 5 keV will necessarily involve effects due to trigger roll-off. The emphasis in future studies, then, must be on a thorough and highly critical study of the efficiency roll-off, both with simulation and, more importantly, direct measurement. A direct measurement can be done with the use of a high-energy gamma emitter such as ^{60}Co (1.17 MeV and 1.33 MeV) or ^{208}Tl (2.61 MeV). The Compton scatter spectrum of such sources is virtually flat, extending to zero energy [29], so any significant deviations from flat behavior can be attributed to varying efficiency. But regardless, the roll-off will likely carry a large systematic uncertainty, and we can therefore conclude that $\sim$5 keV is the lowest energy at which $\mathcal{L}_{\mathrm{eff}}$ can be measured in the absence of an effect from efficiency roll-off. Measurements of $\mathcal{L}_{\mathrm{eff}}$ below this energy can be accepted only after extremely careful scrutiny of the estimated detection efficiency; any significant disagreement between measurement and simulation of this efficiency should be treated with due skepticism.

VI. SUMMARY

The light yield of nuclear recoils in LXe, and its variation with energy, are neatly expressed as the parame-

ter, $\mathcal{L}_{\text{eff}}$. It is rather unfortunate that, at energies most important for dark matter direct detection experiments, measurements of $\mathcal{L}_{\text{eff}}$ largely disagree. Unfortunate, but not unexpected, since such measurements are inundated with many systematic effects that can have large consequences for the results. Sections III and IV exposed the most relevant of these systematic effects and showed not only that detection efficiency roll-off is the largest contributor to the disagreement between results, but exactly how this effect has produced the observed disparity.

When results are ignored that have been plagued by efficiency roll-off, the apparent disagreement disappears. Seen in this way (Figure 6), the available data are all consistent with a featureless $\mathcal{L}_{\text{eff}}$, falling slowly from ~ 0.2 at 100 keV to ~ 0.12 at 5 keV. Coincidentally, this behavior is very well approximated by the best fit from section II.

[1] J. Angle et al. (XENON10 Collaboration), Phys. Rev. Lett. **100**, 021303 (2008), 0706.0039.

[2] V. N. Lebedenko et al. (UK Dark Matter Collaboration), Phys. Rev. **D80**, 052010 (2009), 0812.1150.

[3] E. Aprile et al. (XENON100) (2010), 1005.0380.

[4] G. Bertone, D. Hooper, and J. Silk, Phys. Rept. **405**, 279 (2005), hep-ph/0404175.

[5] O. Cheshnovsky et al., Chem. Phys. Lett. **15**, 475 (1972).

[6] J. Lindhard et al., Mat. Fys. Medd. Dan. Vid. Selsk. **33**, 10 (1963).

[7] A. Hitachi, Astropart. Phys. **24**, 247 (2005).

[8] T. Doke, A. Hitachi, J. Kikuchi, K. Masuda, H. Okada, and E. Shibamura, Jpn. J. Appl. Phys. **41**, 1538 (2002).

[9] A. Manalaysay et al., Rev. Sci. Instrum. **81**, 073303 (2010), 0908.0616.

[10] C. Dahl, Ph.D. thesis, Princeton University, Princeton, New Jersey (2009).

[11] F. Arneodo et al., Nucl. Instrum. Meth. **A449**, 147 (2000).

[12] R. Bernabei et al., EPJdirect C **11**, 1 (2001).

[13] D. Akimov et al., Phys. Lett. **B524**, 245 (2002), hep-ex/0106042.

[14] E. Aprile et al., Phys. Rev. **D72**, 072006 (2005), astro-ph/0503621.

[15] V. Chepel et al., Astropart. Phys. **26**, 58 (2006).

[16] E. Aprile et al., Phys. Rev. **C79**, 045807 (2009), 0810.0274.

[17] A. Manzur et al., Phys. Rev. **C81**, 025808 (2010), 0909.1063.

[18] P. Sorensen et al., Nucl. Instrum. Meth. **A601**, 339 (2009), 0807.0459.

[19] G. Cowan, *Statistical Data Analysis*, Oxford Science Publications (Oxford University Press, Inc., New York, 1998).

[20] A. Hitachi, J. Phys. Conf. Ser. **65**, 012013 (2007).

[21] A. Mangiarotti et al., Nucl. Instrum. Meth. **A580**, 114 (2007), physics/0610286.

[22] D. L. Chichester, J. D. Simpson, and M. Lemchak, J. Radiol. Nucl. Chem. **271**, 629 (2007).

[23] J. D. Simpson (2010), private communication.

[24] A. Manzur (2009, 2010), private communication.

[25] F. Neves et al., IEEE Trans. Nucl. Sci. **52**, 2793 (2005), physics/0501036.

[26] A. Manalaysay, Ph.D. thesis, University of Florida, Gainesville, Florida (2009).

[27] P. Sorensen (XENON10 Collaboration) (2010), presented at the 2010 Light Dark matter Workshop, UC Davis, URL http://particle.physics.ucdavis.edu/seminars/data/media/2010/apr/sorensen.pdf.

[28] P. Sorensen (2010), private communication.

[29] G. F. Knoll, *Radiation Detection and Measurement* (John Wiley Sons; New York, USA, 2000), 3rd ed.